Foundations of Computer Security

Lecture 32: Entropy

Dr. Bill Young
Department of Computer Sciences
University of Texas at Austin

An Encoding Example

Recall the following language:

Example 1: Assume a two-sided coin. Symbols in the language are "H" (heads) and "T" (tails). A string describes the result of repeatedly flipping the coin. E.g., "HHTHTTTHTHTTHTT..."

Suppose you know further that it's a *fair* coin.

The following is the naïve encoding for this language. *Is there a better encoding? How would you know?*

Result	Prob	Code
H	1/2	0
T	1/2	1

Entropy

Claim: If you know the probabilities of symbols in the language, you can compute *the average information content of a symbol in the language.*

For our fair coin example, the computed answer is 1 bit per symbol.

Definition: The *entropy* of a language is a measure of the information content of an average symbol in the language.

Computing Entropy

Entropy H is computed as follows. If p_i is the probability of the ith symbol in the language, then

$$h = -(\sum_i p_i \log_2 p_i)$$

From now on, all log's are base 2.

Example: Consider our fair coin example. Find the entropy.

Result	Prob	Code
H	1/2	0
T	1/2	1

Solution: There are two symbols, each with probability 1/2, so:

$$h = -(1/2 \times \log 1/2 + 1/2 \times \log 1/2) = 1$$

What Does It Mean?

What does it mean to say that the entropy of this language is 1?

1. On average, there is one bit of information in each symbol of the language.
2. It is impossible to find an encoding that uses less than one bit per symbol, on average.
3. Any encoding that uses one bit per symbol, on average, is optimal.

Therefore, for this example, the naïve encoding *is* the optimal encoding.

Entropy Special Case

Whenever you have n symbols, all equally probable, the probability of any of them is $1/n$. The language has entropy:

$$h = -(\log 1/n) = \log n$$

For example, a fair die with six sides has entropy:

$$h = -(\log 1/6) = \log 6 \approx 2.58$$

Hence, it requires 2.58 bits, on average, to transmit the result of a roll.

Another Example

Suppose we have an unbalanced coin that is three times more likely to yield a head than a tail. *What is the entropy of this language?*

Solution: Again there are two possible outcomes:

Result	Prob
H	3/4
T	1/4

The entropy is computed as follows:

$$h = -(3/4 \times \log 3/4 + 1/4 \times \log 1/4) \approx 0.811$$

This says that it's theoretically impossible to encode this language using less than 0.811 bits (on average) to transmit the result of each toss of the coin.

Lessons

- Given the probabilities of symbols in a language, you can compute its *entropy*.
- Entropy measures the average information content of symbols in the language.
- Entropy sets a lower limit on encoding efficiency.

Next lecture: Entropy II